nucleotide sequence at the 3' – end of each mRNA preceding the poly-A (along with restriction site); and (3) allowing for more stringent annealing temperature (e.g., 50°C), thereby reducing potential mispriming during first strand synthesis. Following first strand synthesis, ds cDNA 5' was synthesized using a cocktail of requisite enzymes (DNA Polymerase I, RNase H, and *E. coli* DNA ligase), for example, according to the method of Gubler and Hoffman (Life Technologies Instruction and Technical manuals).

[0051] Figure 4 shows the identification of mRNA in an HEK 293 test sample from cDNA fragments. The identity of the specific mRNAs has been established from combining the specific VNNNN identimer used in the experiment (5' TTTTTTTTTTTTTTTTTTTTGGTTT 3' (SEQ. ID NO. 5)), the specific restriction enzyme employed (NlaIII, cut site is 5' CATG 3'), and the size of the fragments from the gel. Using this information, a search of publicly available mRNA and DNA sequence databases produced the results for the samples shown in Table 1:

REMARKS

The preliminary amendment amends the specification to include a Sequence Listing, to include references to the sequences of the Sequence Listing in the body of the specification and to correct typographical errors. In compliance with 37 CFR 1.821(g), Applicants also state that the Sequence Listing does not add new matter. Further, no new matter is added through the amendments to specification.

A computer readable form of the sequence listing pursuant to 37 C.F.R. 1.821(e) and a statement under 1.821(f) are also included herewith. It is respectfully submitted that the foregoing amendments place the application in compliance with 37 C.F.R. 1.821 through 1.825. Therefore, Applicants respectfully request entry of the amendments and examination on the merits.

It is believed that any fees due with respect to this paper have already been identified in a transmittal that accompanies this paper. However, if any additional fees are required in connection with the filing of this paper, and such fees have not been identified in the accompanying transmittal, permission is given to charge account number 18-0013 in the name of Rader, Fishman & Grauer PLLC.

Respectfully submitted,

Date: July 23, 2002

Customer # 010291

James F. Kamp, Reg. No. 41,882
Christopher J. Voci, Reg. No. 45,184
RADER, FISHMAN & GRAUER PLLC
39533 Woodward Avenue, Suite 140
Bloomfield Hills, MI 48304
(248) 594-0600
Attorneys for Applicants

MARKED UP VERSION OF AMENDED PARAGRAPHS

[0009] The invention comprises compositions and systems to identify the relative expression level of any or all eukaryotic mRNAs in one or more samples. The invention comprises, without limitation, one or more mRNA specific primers for use in reverse transcription that themselves comprises an oligo-dT nucleotide sequence (at the 5' end) linked to a nucleotide sequence (at the 3'end) where the nucleotide immediately adjacent to the oligo-dT segment is not a T. This sequence can be written (from 5' to 3' end) as Tn-VNx, where n = any integer of 8 or greater describing how many T nucleotides are present; V = nucleotides A, G, or C; each N = nucleotides A, G, C, or T, and x = any integer 3 or greater that describes how many N nucleotides are present (SEQ. ID NO. 1). (For purposes of the invention, the designation "d", or "deoxy", shall also include the "nondeoxy" form where appropriate as known by those of ordinary skill). The complete primer (oligo-dT region + VNx sequence) of the invention is called an "identimer."

[0019] With this background in mind, the invention comprises an identimer with three or more nucleotides upstream of a poly-T tail, combined with a restriction enzyme that cleaves ds DNA in a sequence-specific fashion, to generate 3-prime end cDNA fragments of expressed genes. The expression level of a given gene is proportional to and correlates with the amount or abundance of the respective 3' cDNA fragment level. Genes (expressed as mRNAs) are identified by combining the known sequence of 3 or more nucleotides immediately adjacent to the poly-A tail (complementary to the Nx base-anchored primer)(SEQ. ID NO. 1), the specific DNA sequence recognized or cut by the restriction enzyme(s) employed, and the size of the 3' fragment. The size of the 3' fragment represents the distance between the Nx ~~N(x)~~ base-anchored priming (poly-adenylation site) and the nearest restriction enzyme cut site. The identity of the mRNA (gene) may be derived by searching an mRNA or DNA database for the nucleotide sequence that matches the N(x)-base priming site, the restriction enzyme cut site, and the distance between the priming site and the cut site. Ambiguous calls are avoided by repeating the protocol with one or more restriction enzymes that recognize or cut a different nucleic acid sequence.

[0020] In one embodiment, without limitation, the invention comprises up to 192 different identimers that represent all combinations of the primer designated as 5' Tn-VNNN 3' (n = an integer of preferably 21 representing the number of T's; V = nucleotides A, G, or C but not T; each N = nucleotide A, G, C, or T)(SEQ. ID NO. 2), which is designed to identify the 4 nucleotides immediately adjacent to the poly-adenylation site in eukaryotic mRNA. The Tn or poly-dT in the identimer is designed to anneal to the poly-A tail in eukaryotic mRNA. In one embodiment, without limitation, more than one set of 192 identimers (the permutations of VNNN = 3 x 4 x 4 x 4) is employed, and each set is used with a single RNA sample. These identimer sets, and thus the samples, are differentiated by adding a distinct, detectable molecular label or marker, by way of one example only, a fluorescent label, to the 5' end of each identimer in a set, with all identimers within a set having a similar 5' marker. The identimer is annealed to mRNA using buffer and temperature conditions that are known in the art for optimal sequence-specific priming, and reverse transcription is carried out. Second-strand synthesis is subsequently carried out to produce double-stranded ("ds") cDNA that is amenable to restriction enzyme cleavage. Enzyme-mediated, sequence-specific cleavage is carried out, resulting in fragmented ds cDNA. For each set where different cleavage enzymes or agents are used, the invention will generate different 3' end fragments for characterization. In this manner, the invention generates and analyzes cDNA fragments that are assayed for size (e.g., mobility in a gel) and amount.

[0026] In some embodiments, the identification of a gene or mRNA utilizes information derived from the identimer sequence, the restriction enzyme recognition sequence(s), and the size of the resulting cDNA fragments. This information is then employed to search an mRNA sequence database to identify the specific genes or mRNAs in the samples under investigation. The data used to search the mRNA database are derived by means of the invention. The mRNA nucleotide sequence of the bases immediately adjacent to the poly-A tail are derived from knowledge of the complementary identimer sequence. For example, if the identimer for a given reaction has the sequence 5'-TTTTTTTTTTTTTTTTTTTTAAAC-3' (SEQ. ID NO. 3), then any mRNAs identified from this reaction will contain the sequence 5'-GTTTAAAAAAAAAAAAAAAAAAA-3' (SEQ. ID NO. 4). Further information is derived from the determination of the length of labeled cDNA fragments and the restriction enzyme employed to generate the fragments. For example, if the first restriction digest of the

identimer reaction above employs the restriction enzyme NlaIII ~~NLAIII~~, which cuts at the sequence 5'-CATG-'3, then a cDNA fragment that is 334 bases in length identifies and mRNA sequence that contains the 5'-CATG-'3 sequence 314 bases from the poly-adenylation site. This take into account the 20 "T" bases on the identimer (i.e. 334 - 20 = 314). If the second restriction digest of the identimer reaction employs the restriction enzyme MboI ~~MBOI~~, which cuts at the sequence 5'-GATC-3', then a cDNA fragment that is 889 bases in length identifies an mRNA sequence that contains the 5'-GATC-5' sequence 869 bases from the poly-adenylation sequence. Using this information to search an appropriate database, one can identify the mRNA as human precerebellin (GI# 180250), which matches the analytical data. If no mRNA is present in the database, then one can employ a similar bioinformatical strategy to predict the identity of the unknown mRNA or approximate the identity of mRNA or gene family involved. Similarly, if the samples are derived from an organism that lacks an adequate mRNA or gene sequence database, the mRNA is identified using the database from a closely related species.

[0034] First and second strand cDNA synthesis. First strand synthesis is performed by means known to those of ordinary skill (using any experimental cell/tissue type) on the total RNA population utilizing a four – base identimer of sequence $NNNVT_{21}$, where each N = A, C, T, or G, and V = A, C, or G but not T (SEQ. ID NO. 2). In practical application, the total number of unique identimer tags theoretically required to span the total estimated mRNA population (in a eukaryotic organism) would be 192 (thus 192 unique subsets). Compared with most differential display protocols, which typically utilize a two – base anchored primer for first strand synthesis, a four – based identimer has advantages by: (1) significantly reducing the complexity of the mRNA pool by a factor of 16 (192/12 = ~ 16), thereby reducing the number of bands displayed per fingerprint (or subset); (2) providing a more accurate prediction of the candidate mRNA(s) of interest through the additional two nucleotide sequence at the 3' – end of each mRNA preceding the poly-A (along with restriction site); and (3) allowing for more stringent annealing temperature (e.g., 50°C), thereby reducing potential mispriming during first strand synthesis. Following first strand synthesis, ds cDNA 5' was synthesized using a cocktail of requisite enzymes (DNA Polymerase I, RNase H, and *E. coli* DNA ligase), for example, according to the method of Gubler and Hoffman (Life Technologies Instruction and Technical manuals).

[0051] Figure 4 shows the identification of mRNA in an HEK 293 test sample from cDNA fragments. The identity of the specific mRNAs has been established from combining the specific VNNNN identimer used in the experiment (5' TTTTTTTTTTTTTTTTTTTTGGTTT 3' (SEQ. ID NO. 5)), the specific restriction enzyme employed (NlaIII, cut site is 5' CATG 3'), and the size of the fragments from the gel. Using this information, a search of publicly available mRNA and DNA sequence databases produced the results for the samples shown in Table 1:

R0131269